Highly Available Atomic Consistency

Ivan Klianev
Transactum Pty Ltd

High Performance Transaction Systems
Asilomar, California, 27-30 September 2015
How We Built a Data System where:

- CAP theorem is *Inapplicable*
- Distributed Transactions *Do Not Block*
- Synchronous Replication brings *No Additional Latency*
- Multi Partition Transactions *Outperform* Local Ones
Clustered Network Links

- CAP Theorem to Apply: A Faulty Network Link Partitions the System
- To Bypass the Theorem: Prevention of System Partitioning is Sufficient

Our Clustering Solution:
- Integrates Redundant Capacity
- Prevents Lost or Delayed Messages
- Guarantees Ordered Delivery
System Architecture

Application Servers

Transaction Managers

Scaled Out Database

Replicas Partition 1

Replicas Partition N

Network Links Cluster
Tx Request / Response

Network Links Cluster
Tx State Replication

Network Links Cluster
Tx Effects Replication

HPTS, Sept. 2015, Asilomar, California, USA

Transactum Pty Ltd, Sydney, Australia
Architecture of Transactions

- 'Writes' executed on All Transaction Managers
- Against their own Main Memory Copy of the database
- Effects of 'Writes' applied on Multiple Replicas
Synchronous Linearizable Replication

Performance Evaluation

TPC-C benchmark New-Order transactions with 10 items on 3 different systems:

- Classic Tx
  - No Replication

- Sync Replication
  - One Data Replica
  - Two Data Replicas

Application Server
Machine under $500

Transaction Management
Machines under $1,500

Data Management
Machines under $500
running MS SQL
Synchronous Linearizable Replication

Result: No Added Latency

- **No Replication**
- **Classic Transactions**
- **Synchronous Replication on One Data Replica**
- **Synchronous Replication on Two Data Replicas**

Graph showing the performance comparison with **Background Flush** and **Periodic Full Checkpoint** for different replication scenarios.

HPTS, Sept. 2015, Asilomar, California, USA

Transactum Pty Ltd, Sydney, Australia
Multi Partition Transactions

Performance Evaluation

2 Partitions, TPC-C benchmark New-Order transactions with 5 items from each Partition

Single Partition Transactions

Two Partition Transactions, Two Replicas per Partition
Multi Partition Transactions
Result: Higher Throughput

One Partition
Classic Transactions
Non-Replicated: 95%
Synchronously Replicated: 100%

One Partition
Our Architecture
Non-Replicated: 100%
Synchronously Replicated: 135%

Two Partitions
Our Architecture
Synchronously Replicated: 135%

HPTS, Sept. 2015, Asilomar, California, USA
Transactum Pty Ltd, Sydney, Australia
Conclusion

The presented Transactional Architecture guarantees

Atomic Consistency of Replicas
High Availability of Transactions
Higher Throughput with Multi Partition Transactions
Thank You

Ivan Klianev
Transactum Pty Ltd